1973-06

Analysis of a random communication network by simulation.

Hung, Nguyen dich

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/16463

ANALYSIS OF A RANDOM COMMUNICATION NETWORK BY SIMULATION

Nguyen dich Hung

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

Analysis of a Random Communication Network

by Simulation

by

Nguyen dich Hung

Thesis Advisor: Shu-Gar Chan

June 1973

Analysis of a Random Communication Network

by Simulation

by

Nguyen dich Hung
Commander, Vietnamese Navy

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 1973

ABSTRACT

In this paper a unified presentation is made on the results of various investigators on the properties of random communication networks. These results are interpreted in such a way that the properties may be determined by using a digital computer with the application of the Monte Carlo method. The computer program is written and tested. Results for some networks are compared with theoretical values.

TABLE OF CONTENTS

ACKNOWLEDGEMENT

The author wishes to express his appreciation to Professor Shu-Gar Chan for the invaluable aid and counsel which he has offered during the preparation of this thesis.

I. RANDOM COMMUNICATION NETWORK

In this paper the results of various investigators on the properties of random communication networks are reviewed. From these results a mathematical model is represented for a general random communication network.

A. THE CONCEPT OF A RANDOM COMMUNICATION NETWORK

A communication network is an aggregate of message centers that attempt to transfer information to one another over a wide variety of channels. The message centers could be fixed as the headquarters of a regional military command, mobile as ships, aircrafts, or satellites. The channels could be a radio link, a telephone line, or highway. In normal conditions, the topology of the network is known with certainty: the positions of ships at sea, the links between the headquarters with different units. Each center knows its relative position with respect to the other centers exactly. If a message is sent from a central command, the authorities know exactly how many and which units will receive it. The network, in this case, is called a deterministic communication network. But during the hostilities, some units could be destroyed by the enemy without even having the chance to send the last word, some radio link could be jammed by the enemy, some telephone lines or some roads could be damaged without having any means to determine the extent of the destruction. Back at the headquarters, the authorities are faced with the uncertainty of the existence of the different units and channels, the effectiveness of an order sent to the battlefield is not as high as before. It is only probable that such unit gets the message. The communication network becomes a highly probabilistic, random netwrok in the event of a war.

One finds the same kind of randomness in many other branches of research; for example, the contagion of a disease among a population, an hereditary defect in a species, the spread of a stimulus in the nervous system, or the spread of information in society, to name a few. All these phenomena are very similar to the random communication network in that they have:

a. an originator which originates a specific message

b. several recipients which convey the received 'message' to the others via a medium (atmosphere, voice, physical contact, etc.)

c. the uncertainty of contact between two individuals.

All these phenomena could be then studied under the same topic of the random communication network. Researches in these phenomena give much insight in the problem of propagation message through a random network.

As it is said, the communication between two message centers A and B could be assumed only with a certain probability. The latter measure is a function of many parameters which are for example, the time t when the message is sent, the time τ since the message is received at station A, the distance d between stations, the electronic equipment of each station, to cite a few. Therefore one could write, in general:

$$\text{Prob (contact A to B)} = p\ (A, B, t, \tau, d, \ldots).$$

Due to the complexity of the problem, without reasonable assumptions about the parameters to simplify the problem, one could not solve it. Following is the brief survey of the contributions of various workers. Each of them has solved partially the problem with different assumptions

about the parameters which determine the probability of contact between two stations, two individuals, whatever is the case.

B. THEORETICAL DEVELOPMENT

Bailey [1] studied the spread of an epidemic in an homogeneous population with size n where the probability of one new infection taking place in the interval dt is given by

Prob (one infection in dt) = $y\,(n-y+1)dt$,

where y = the number of susceptibles.

From the above equation, the mean m(t) of the number of infected people and the epidemic curve, (i.e., the rate of change with respect to time of the mean of the number of infected people) $\frac{dm(t)}{dt}$ are obtained to be

$$m(t) = e^{-10t}(810t - 234) + e^{-18t}(4410t - 902) + \ldots$$

$$z = \frac{dm(t)}{dt} = e^{-10t}(8100t - 3156) + e^{-18t}(79{,}380t - 20650) + \ldots$$

Landau and Rapoport [2] assumed the probability of contact between any two individuals is the same for any pair and the probability of transmission of the disease depends both on the time t of the whole process and also on τ, the time since the particular affected individual acquired the disease. Thus,

Prob (one infection) = $p(t,\tau)$

The differential equations giving the number of individuals who become affected and the rate are:

$$\frac{dx(t)}{dt} = \alpha(N - x)\left[x_0 p(t,t) + \int_0^t \frac{dx(\lambda)}{d\lambda} p(t, t-\lambda)\, d\lambda\right] \quad (1.1)$$

$$\frac{dz(t)}{dt} = \frac{1}{N}\frac{dx(t)}{dt} = \alpha N(1 - z)\left[\frac{x_0}{N} = p(t,t) \int_0^t \frac{dz}{d} p(t, t-\lambda)\, d\lambda\right] \quad (1.2)$$

where N = number of individuals in the population

$x(t)$ = total of individuals who have become affected up to time t

x_0 = number of individuals who become affected at t = o, the initial time of the process

α =frequency of contacts, which is assumed constant. . More precisely $\alpha x(t)[N - x(t)]$ is the number of contacts per unit time between affected and unaffected individuals.

This model is more realistic than that of Bailey because it is true that in the case of a disease epidemic, the infectiousness of the pathogenic virus may increase or decrease with the time of the process and the time since it infects the body of the individual.

The equations (1.1) and (1.2) are difficult to solve in the general case, therefore the investigators made some assumptions and solve them for the following cases:

a. The probability $p(t,\tau)$ is a function of t alone

b. The probability $p(t,\tau)$ equals $e^{-k\tau}$

c. The probability $p(t,\tau)$ is function of τ only

d. The probability $p(t)$ is constant in a finite interval

e. There is a lag in the transmission of contagion.

Landahl [3] treated the spread of some information as a flow of 'particles' which execute random motions over a population of individuals and which may multiply or disappear. Equations are derived for the

density of these particles, and the number of individuals through which the 'particles' have passed, is calculated. The results are applied to the following cases:

a. Uniform spatial distribution with multiplication factor decreasing with time because of loss of interest.

b. Multiplication factor is constant, but the rate of spread decreases with multiple hearings.

c. One dimensional region with a small starting region with or without an absorbing barrier. An absorbing barrier corresponds to the case where the individuals listen to the information but do not repeat it.

d. Two-dimensional region with absorbing barrier.

e. Continuous sources of information within a small region in one dimension.

f. Uniform spatial distribution in which individuals do not respond to more than one hearing.

Prihar [4] studied a mobile communication network where the transmitting and receiving stations are in motion. He used the classical Kinetic Theory of gases to compute the expected total number of vehicles contacted per unit time in two cases:

a. The search for contact is continuous and the antenna is omnidirectional.

b. The search is intermittent and the antenna is directional.

Finally, Mattei [5] assumed the probability of contact between any pair of stations is constant throughout the network.

Prob (contact between any pair) = p = constant.

Several formulas were derived in closed forms.

The mean number of contacted nodes at the 1st step is

$$\bar{Z}(1) = (n - 1)p.$$

The mean number of contacted nodes at the 2nd step is

$$\bar{Z}(2) = (n - 1)q\ [1 - (1 - p^2)^{n-2}].$$

The marginal probability of the number of new nodes at each step is

$$\text{Prob}[Z(k) = z(k)] = \sum_{z(1)=0}^{y(0)} \ \sum_{z(2)=0}^{y(1)} \cdots \sum_{z(k-1)=0}^{y(k-2)}$$

$$\left\{\binom{y(0)}{z(1)} \ [1 - q^{z(0)}]^{z(1)} \ [q^{z(0)}]^{y(1)}\right\}$$

$$\left\{\binom{y(1)}{z(2)} \ [1 - q^{z(1)}]^{z(2)} \ [q^{z(1)}]^{y(2)}\right\}$$

$$\left\{\binom{y(k-1)}{z(k)} \ [1 - q^{z(k-1)}]^{z(k)} \ [q^{z(k-1)}]^{y(k)}\right\}$$

$$y(k) = n - \sum_{i=0}^{k} z(i) \qquad q = 1-p \tag{1.3}$$

where n = number of nodes in the network

k = order of step

C. THE PROPOSED MATHEMATICAL MODEL

Although with the model of Mattei, one could derive many quantities of the network in closed and neat forms, the model is not realistic in real life as pointed out in the beginning of this paper. The probability of contact between any pair of stations is far from constant. It

depends on their distances to the center of attack, the size of the electronic equipments, the geographical positions. In the event of a war, the capital, the missile silos, the industrial centers are more vulnerable than a submarine under the sea or a communication satellite in orbit.

In this paper, a model, in which the probability of contact between any pair of stations could be varied at will, could be assigned to each link, once all circumstances have been taken into account, is proposed and solved. The solution is obtained by simulation on the digital computer. The computer program allows the user to assign the probability of contact between all pairs of stations.

II. SIMULATION ON A DIGITAL COMPUTER

The aim of this chapter is to obtain, by simulation on the digital computer using the Monte Carlo method, the desired quantities of a general random communication network, in which the link probabilities could be all distinct.

A. THE MONTE CARLO METHOD

The Monte Carlo method consists of solving various problems by means of the construction of some random process for each such problem, with the parameters of the process equal to the required quantities of the problem. These quantities are then determined by means of the observations of the random process and the computation of its statistical characteristics which are approximately equal to the required parameters.

For example, the required quantity x might be the mathematical expectation $E(\eta)$ of a certain random . The Monte Carlo method for determining the approximate value of the quantity x consists of an N-fold sampling of the value of the variable η in a series of independent tests, (that is, $\eta 1$, $\eta 2$, ...) and the computation of their mean value

$$\eta = \frac{\eta_1 + \eta_2 + \cdots + \eta_N}{N}$$

If the number of tests is large, one has

$$\bar{\eta} \simeq E(\eta) = x.$$

It is appropriate to ask the question why one does not try to derive an analytical expression for the quantity x from the properties of the problem instead of designing a random process which imitates the reality. The answer is that many problems in the physical world are so complex that they do not have closed formula for the desired quantities.

Suppose, for example, that the required quantity x is the definite integral of a function f(y) taken over the interval (a, b), that is

$$x = \int_a^b f(y)\,dy.$$

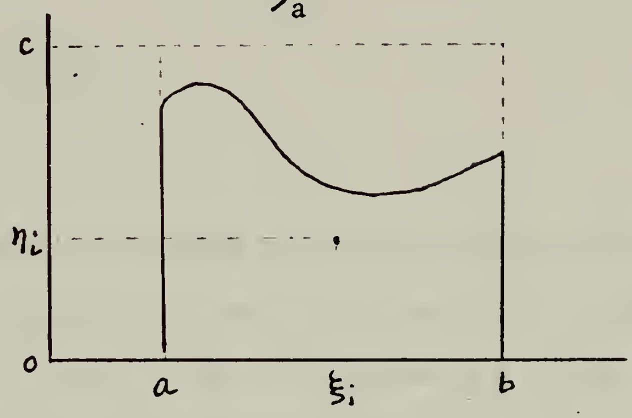


Fig. 2.1. Illustrating equation (2.2).

If the expression of f(y) is not simple (or even worse, if the function f(y) does not have a closed formula but is known only point by point), x could not be evaluated analytically. The Monte Carlo method may be used as will be described next.

Suppose we have a device called a random number generator which generates two random variables ξ and η uniformly distributed over the intervals (a, b) and (o, c), respectively. The Compound probability density function is given by

$$P_{\xi,\eta}\ (y,z) = \begin{cases} \dfrac{1}{c(b-a)} & \text{if } a<y<b \text{ and } 0<z<c \\ 0 & \text{otherwise} \end{cases}$$

The pair (ξ,η) will be generated N times and each time, the condition

$$f(\xi_i) \lessgtr \eta_i \qquad (2.1)$$

will be tested. If condition (2.1) holds, the point (ξ_i, η_i) in Fig. 2.1 is inside or on the boundary of the computed area. Let N' be the number of times when this condition holds in N tests, we have

$$x = \int_a^b f(y)dy = c(b-a) \lim_{N\to\infty} \left(\frac{N'}{N}\right) \qquad (2.2)$$

In order to obtain good results, the number of tests must be large, say, in the range between 1,000 and 100,000. This is only possible with the aid of digital computers.

The problem of a random communication network where the probability of survival of each link is assumed constant has been solved adequately by Mattei [5]. If one changes this assumption and lets the said probability vary to fit certain conditions in reality, the problem becomes untractable, then the only feasible solution is to use a numerical method such as the Monte Carlo technique. The required quantities are: the average number of contracted nodes, the average number of newly contacted nodes at each step, the terminal reliability, the probability distribution of new nodes at each step. The method for obtaining these quantities will be described in the following section.

B. THE FLOWCHART OF THE SIMULATION PROCESS

Suppose that one has a network consisted of n stations connected by a certain number of links, it is a deterministic network in the sense that, if a message is sent by a station to adjacent stations which relay it to the others, one knows exactly how many and what stations have received it. Suppose a disaster, which could be a nuclear attack, an earthquake, etc., happens to the network. Before any investigation of the damage could be made, one does not know which links have survived. If a message is sent, it is difficult to acertain how many and which stations will get the message. The best one could do is to assign to

each link a certain probability of survival. Let P_{ij} be the probability of communication between station i and station j, the probabilities between all pairs of stations are best described by an n x n matrix as

$$\underset{\sim}{P} = \begin{pmatrix} 0 & P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & 0 & P_{23} & \cdots & P_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ P_{n1} & P_{n2} & \cdots & P_{n,n-1} & 0 \end{pmatrix}$$

where P is called the matrix of probabilities. Since one is not interested in the case where a station sends back to iteself a received message, all diagonal elements of $\underset{\sim}{P}$ are zero. From a set of n stations, one could form $2^{n(n-1)}$ networks. Fig. 2.2 shows two of them for n = 5.

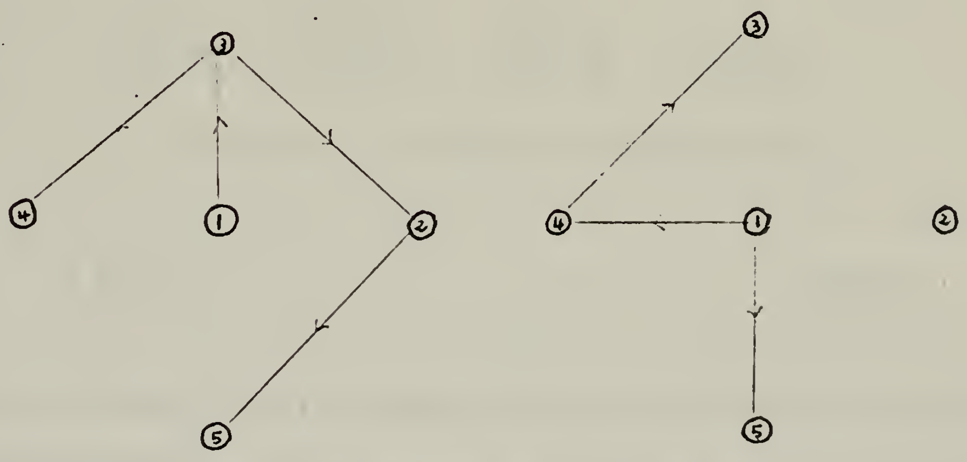
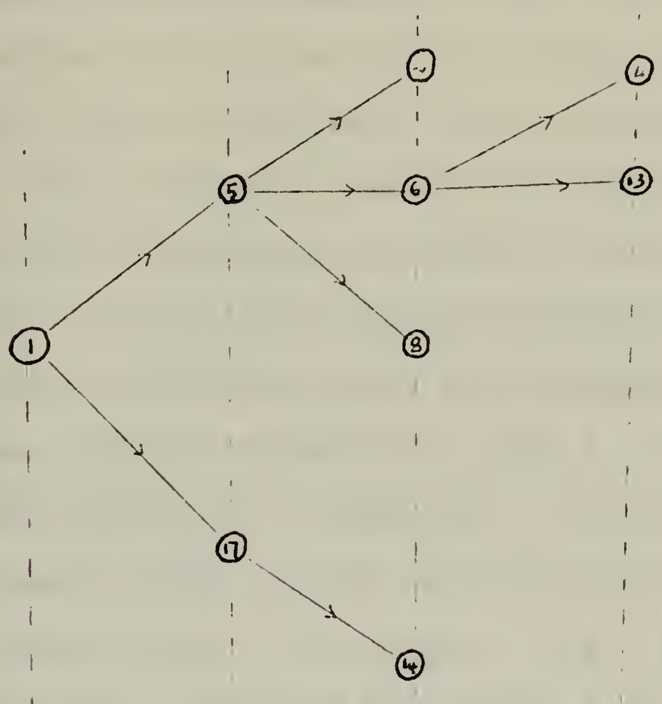


Fig. 2.2. Samples of random network

If the number of stations increases, say n = 21 stations, the population of possible networks is $2^{(20 \times 21)} \simeq 2.7 \times 10^{126}$, which is large. Suppose that one investigates the propagation of message via the originator which is the station 1. Starting with this station, one generates a random variable r uniformly distributed between 0 and 1, and check the condition



step 0 step 1 step 2 step 3
$z(0)=1$ $z(1)=2$ $z(2)=4$ $z(3)=2$ $z(4)=0$

Fig. 2.3. A tree of propagation

$$r \leqslant P_{1,2}. \tag{2.3}$$

If this condition holds, there is a contact between station 1 and station 2 and one increases the random variable z(1), which represents the number of newly contacted nodes at

step 1, by one unit. One generates again a random number r, tests condition (2.3) for $P_{1,3}$ and increases z(1) by one unit if condition (2.3) holds, and so on until one arrives at the last station, which is station 21. One repeats the same procedure with each station of z(1) relative to the remaining network, the number of stations contacted at this second step is designated z(2). One repeats the same procedure with each station of z(2). The procedure stops at the step when one could not contact any more stations.

As an example, consider the network in Fig 2.3, stations 5 and 17 have been contacted at step 1 so that z(1) = 2. The stations which have received the message are discarded from the network, these are stations 1, 5, 17. Next, to each station contacted at step 1, the same procedure made for station 1, is repeated. In the example, stations 2, 6, 8 and station 14 have been contacted respectively by station 5 and station 17 at step 2, z(2) = 4. Stations 4 and 13 have been contacted by station 6 at step 3, z(3) = 2. The propagation of the message is over at step 4 because stations 4 and 13 cannot carry the message any further.

The tree in Fig. 2.3 summarizes the process in various steps. At any time the computer needs to keep the information of only one tree, which saves a lot of memory storage. To start the next sampling, one just has to erase all information of the first sampling run except, of course, the results z(1), z(2), ... The general procedure for the second sampling is the same as for the first.

After K samplings, the average number of new nodes at each step is given by

$$\overline{Z}(1) = \frac{Z_1(1) + Z_2(1) + Z_3(1) + \ldots + Z_k(1)}{K}$$

$$\overline{Z}(2) = \frac{Z_1(2) + Z_2(2) + Z_3(2) + \ldots + Z_k(2)}{K}$$

$$\begin{array}{ccccc} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & & \end{array}$$

$$Z(KMAX) = \frac{Z_1(KMAX) + Z_2(KMAX) + \ldots + Z_k(KMAX)}{K}.$$

where KMAX = the highest order of steps in the tree.

The average number of contacted nodes is

$$\overline{X} = \overline{Z}(1) + \overline{Z}(2) + \ldots + \overline{Z}(KMAX).$$

$$\overline{X} = \frac{\begin{array}{l} Z_1(1) + Z_1(2) + \ldots + \overline{Z}_1(KMAX) \\ + Z_2(1) + Z_2(2) + \ldots + Z_2(KMAX) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ + Z_k(1) + Z_k(2) + \ldots + Z_k(KMAX) \end{array}}{K}$$

$$= \overline{X}(1) + \overline{X}(2) + \ldots + \overline{X}(KMAX)$$

where $\overline{X}(i)$ is the average number of contacted nodes up to step i.

The Kth terminal reliability is commonly defined as

$$r(K) = \frac{\overline{Z}(K)}{n-1}$$

and the weak connectivity is defined as

$$\gamma = \frac{\overline{X}(KMAX)}{n}$$

<u>Definition</u>. In this paper the link density of the network is defined as the ratio of the number of links between stations to the number of emitting stations.

During the process of simulation, the probability distribution of new nodes at each step is also calculated and displayed in an n x n matrix whose rows represent the number of steps and columns represent the number of new nodes as follows:

\# of contacted stations

$$\begin{array}{c c}
\begin{matrix} 0 & 1 & 2 & 3 & . & . & . & n \end{matrix} & \\
\left(\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix}\right) & \begin{matrix} 0 \\ 1 \\ 2 \\ . \\ . \\ . \\ n \end{matrix} \quad \text{order of step}
\end{array}$$

Fig. 2.4. Probability distribution of the number of contacted stations at each step.

A flowchart of the simulation is drawn in detail in Fig. 2.5. The list of the variables used in the computer program is described as follows.

IRMAIN = vector whose components have the value 1 or 0. If the ith component is 1, it means station i has yet been contacted, 0 means the station ith has been contacted.

NEW = vector whose components have values 1 or 0. If the ith component is 1, it means the station i has just been contacted during the preceeding step.

ICOPY = vector with 1 or 0 as components, if the ith component is 1, it means the station i has been contacted during the current step. Before to go to the next step, one has to make; NEW = ICOPY.

ZVEC = vector whose ith component represents the number of newly contacted stations at step (i - 1).

XVEC = vector whose ith component represents the average number of contacted nodes at step (i - 1).

GVEC = vector whose ith component represents the (i - 1)th terminal reliability.

RELIAB = vector whose ith component represents the (i - 1) terminal reliability.

= vector whose ith component represents the probability distribution of new nodes.

RNGE = vector whose components determine the x, y ranges of the plot in the subroutine UTPLOT to plot the average number of new nodes.

PRDIST = square matrix whose (i, j) component represents the probability that the number of contacted nodes is j at step (i - 1).

PROB = square matrix whose (i, j) component represents the probability of communication from station i to station j.

RLIAB = square matrix whose ith row represents the terminal reliability for the ith case (different probabilities, different standard deviations, etc...).

XXVEC = rectangular matrix whose ith row represents the average message propagation for the ith case (different probabilities, different standard deviations, etc...).

KSTEP = integer which represents the order of the step.

KHECK = integer whose value is 1 or 0 used to initialize

KHECK1 = ICOPY to (0) or to compute the number of emitting stations.

SUMND = number of contacted nodes up to step K.

ZNODES = number of contacted nodes during step K.

KMAX = highest order of step attained during the process of simulation.

KKMAX = KSTEP + 2 used in the computation of PRDIST minor variable.

CONECT = represents the weak connectivity of the network.

NDEMIT = the number of emitting stations during one simulation.

LABEL = used in the subroutine DRAW.

MC = integer used in the subroutine DRAW.

IX = integer used in the subroutine OVFLOW and RANDOM.

A = the average number of contacted stations at the 1st step in the model of Mattei or defined as n/A, standard deviation in the general model.

S = the standard deviation of destruction in the general model, defined as: S = n/A.

X and Z = Normalized variables, defined as X = I/S and Z = J/S.

K = the number of tests in the process of simulation.

P and PX = outputs of the subroutine NDTR which are areas under the normal curve.

The following subroutines of the NPGS computer center library are used in the program.

UTPLOT = to plot the average number of new stations curve.

DRAW = to plot the probability distribution of connected stations, the terminal reliability and the average number of contacted stations.

OVFLOW, RANDOM = used to get a random variable (0,1).

NDTR = to obtain the area under the normal curve as probability of communication between two stations.

To investigate the general model where the destruction is assumed normally distributed around the center of attack, block(2) has to be replaced by block(1).

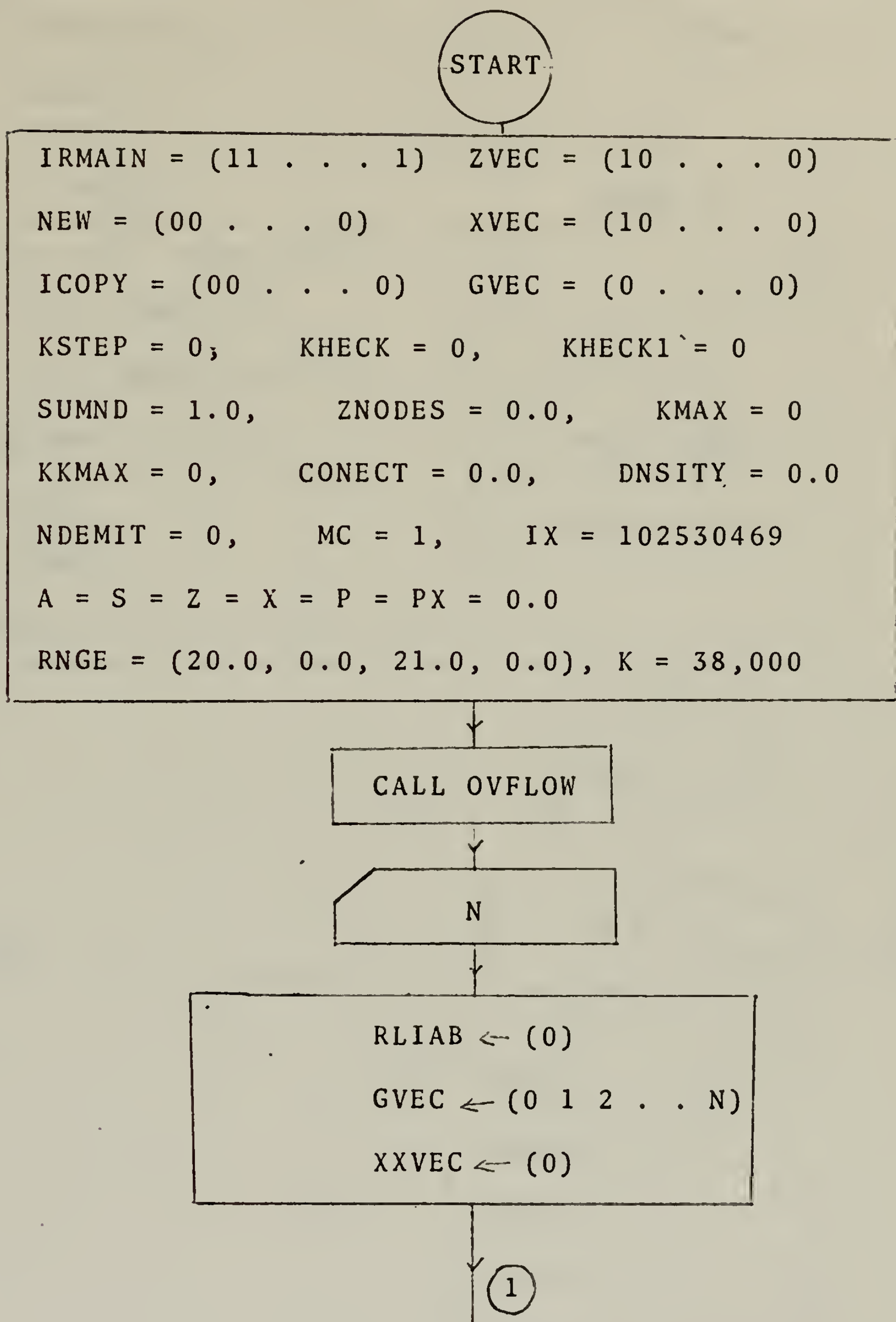


Fig. 2.5. The Flowchart of the simulation.

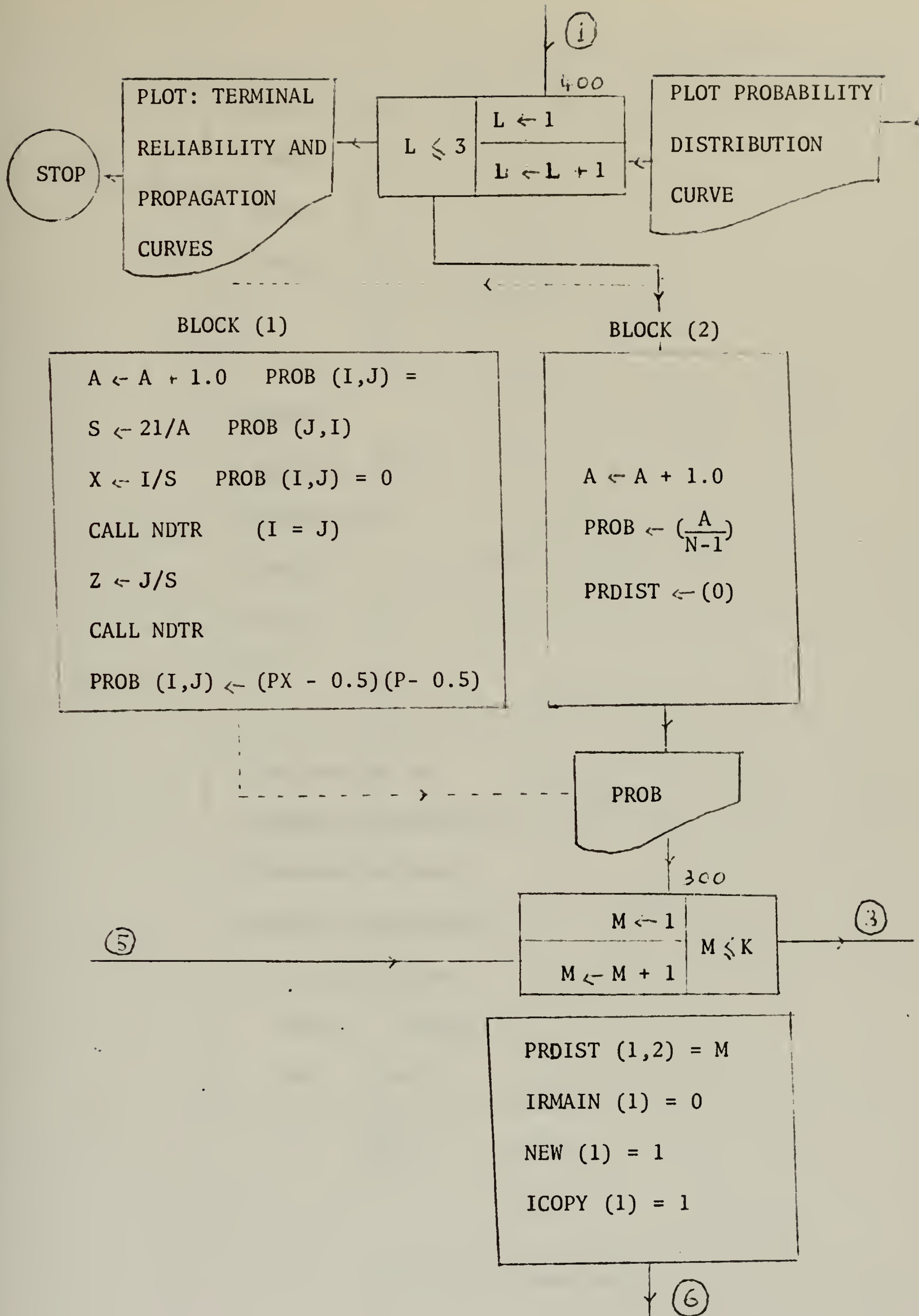


Fig. 2.5. continued.

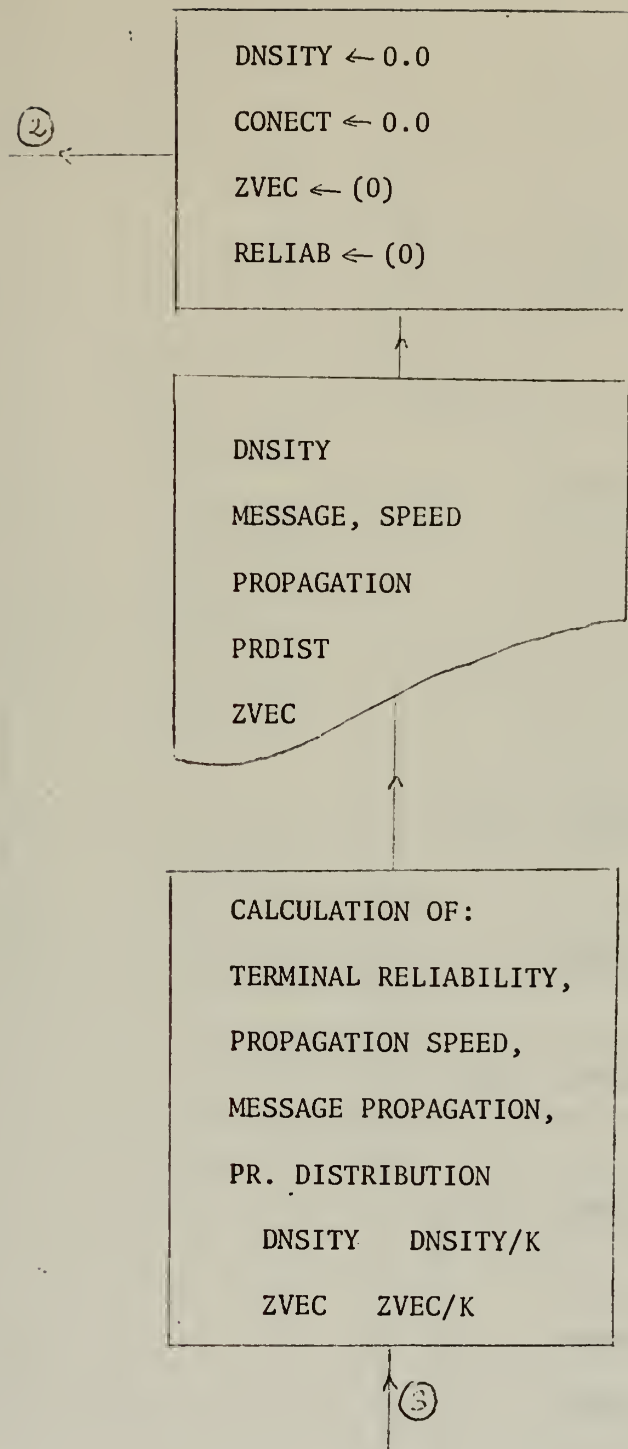


Fig. 2.5. continued.

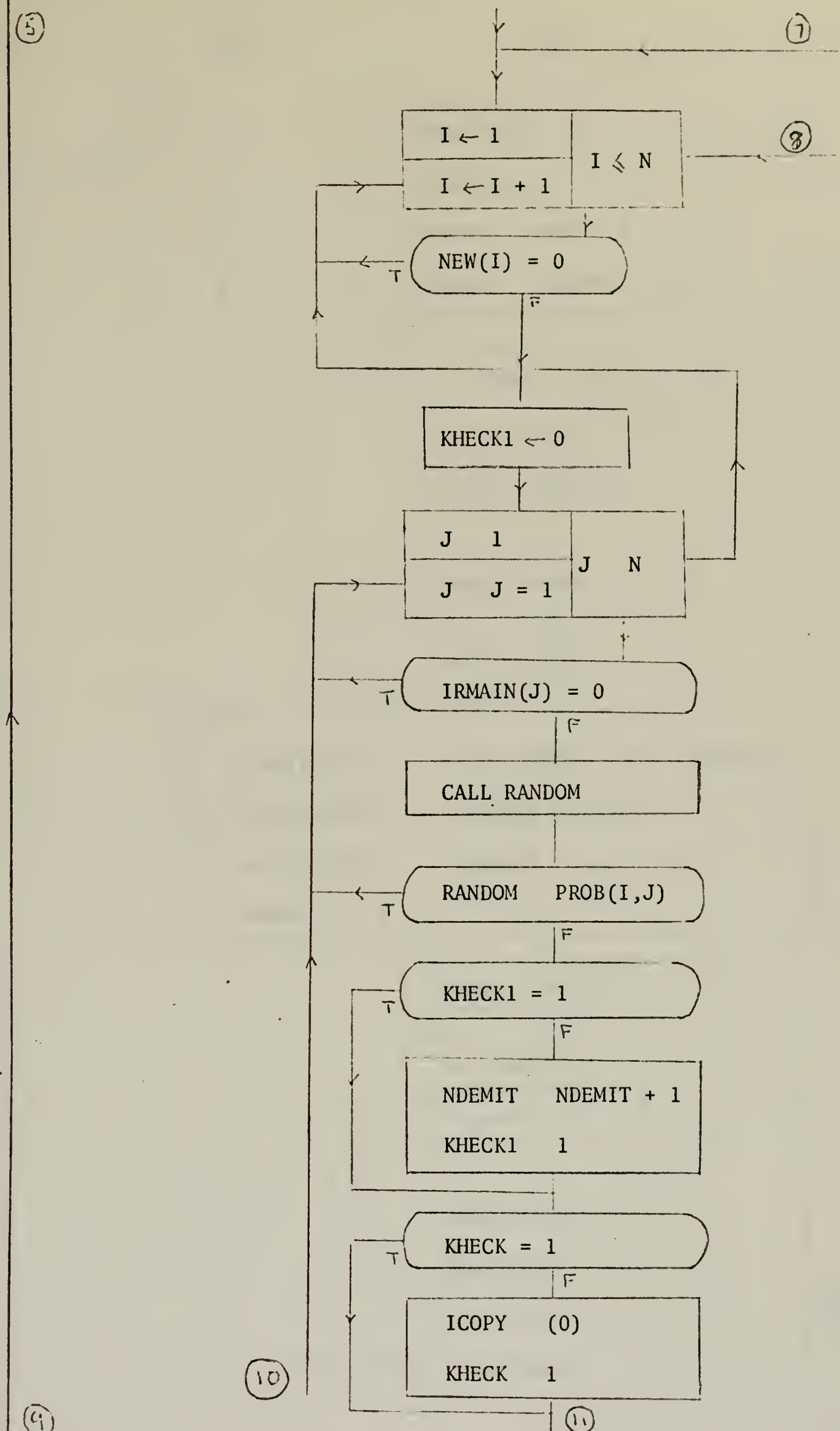


Fig. 2.5. continued.

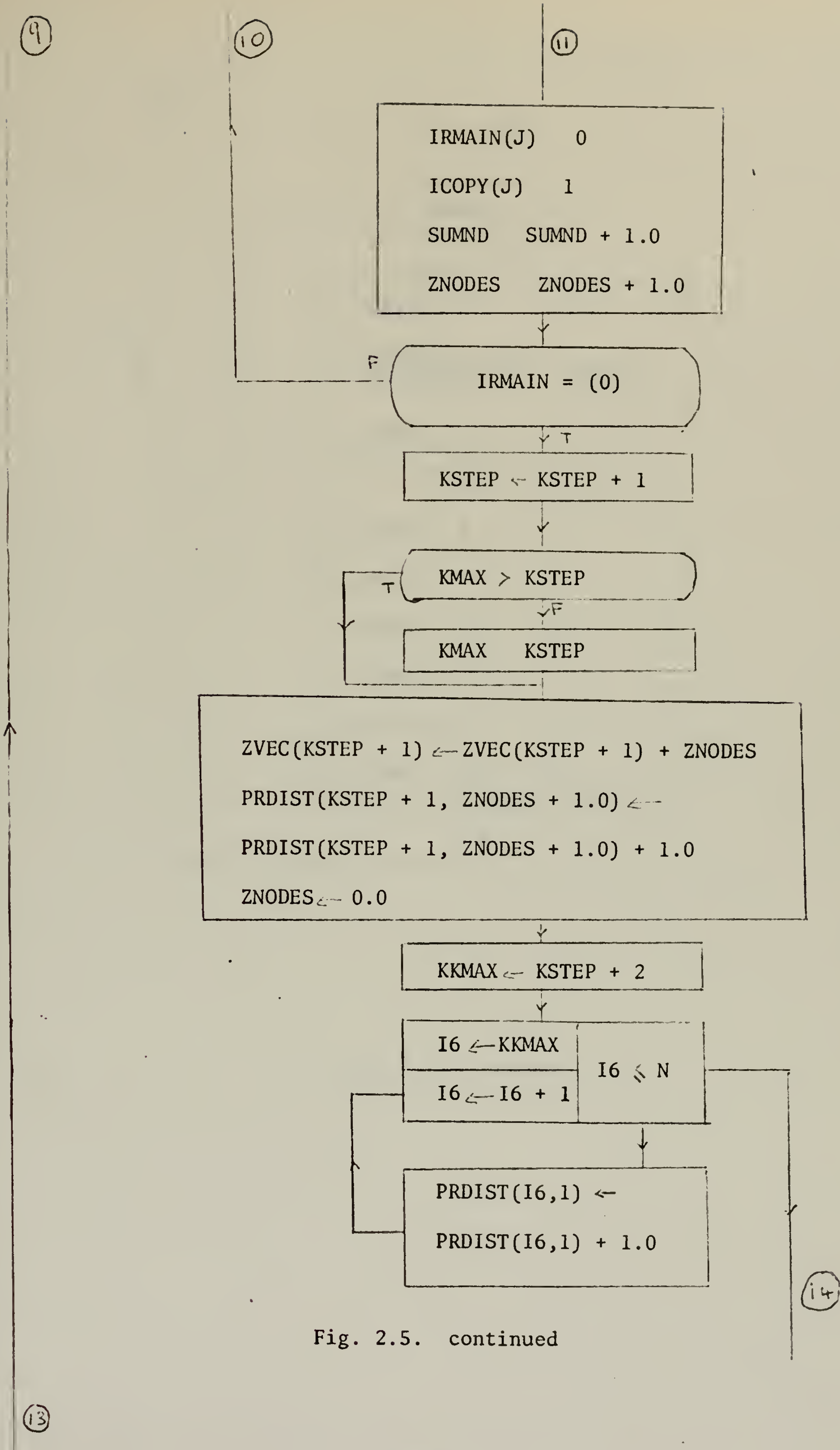


Fig. 2.5. continued

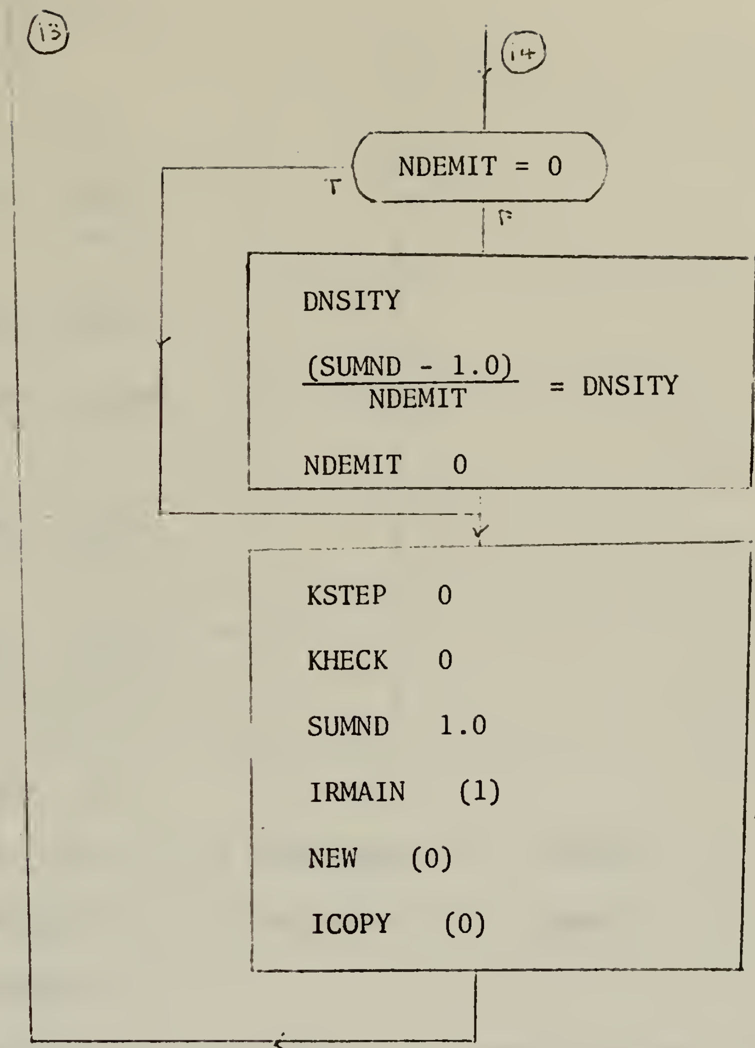


Fig. 2.5. continued.

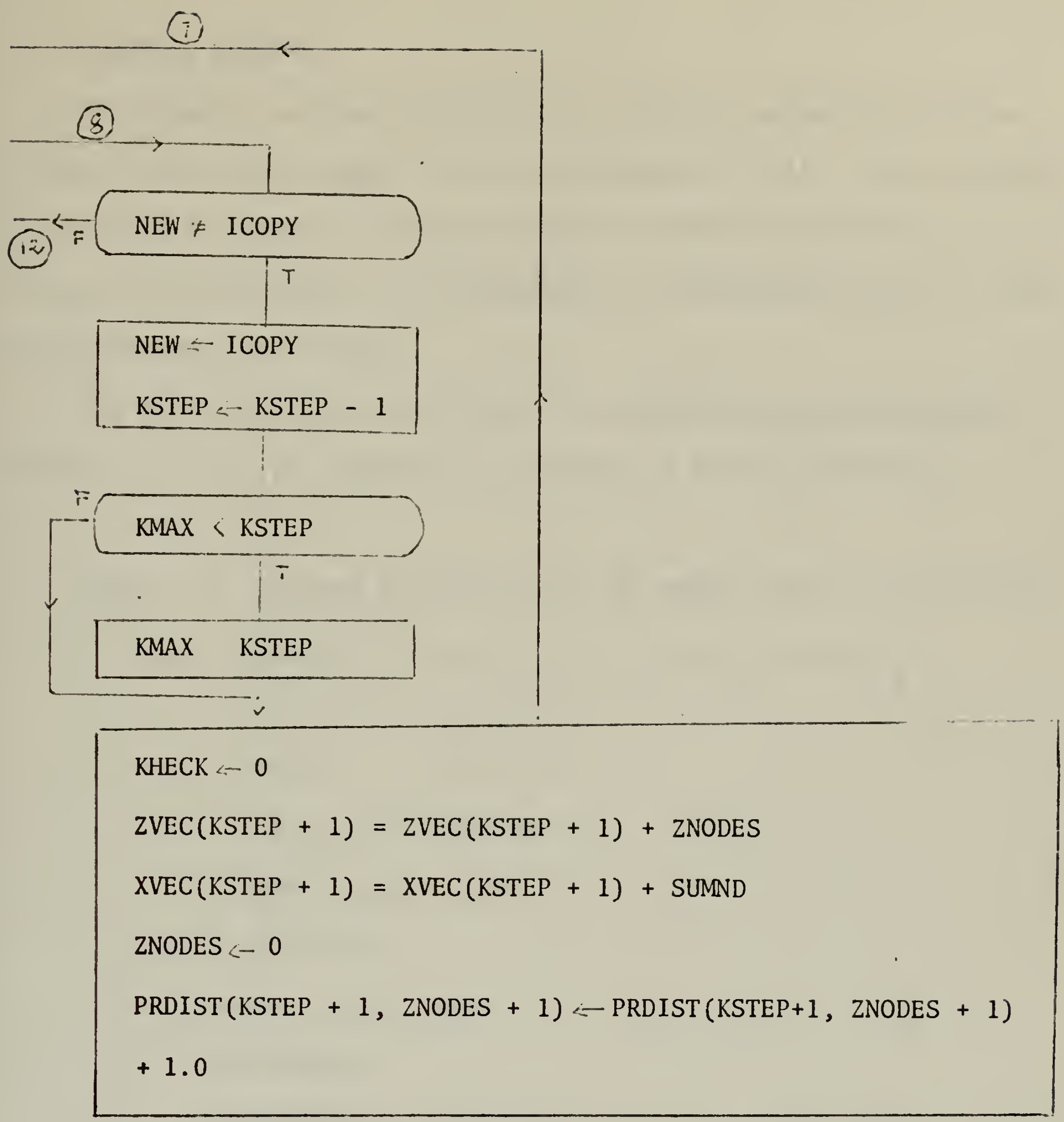


Fig. 2.5. continued

C. COMPUTER PROGRAM

The computer program is written to analyze a network of 21 stations, however this number could be increased at will. First one has initialize the matrix of probabilities of communication between stations to desired values. The probability distribution function could be discrete or continuous.

For example, in the case where the probability of communication between any pair of stations is constant, it suffices to write

Table 2.1. Fortran statements for the model where p is constant

```
      DO 16       I = 1, N
      DO 15       J = 1, N
      IF (I. EQ. J) GO TO 14
      PROB (I,J) = A/(N - 1)
      GO TO 15
   14 PROB (I,J) = 0.0
   15 CONTINUE
   16 CONTINUE
```

If one wishes to analyze a network where the destruction is normally distributed around the center of a nuclear attack, it suffices to replace table 2.1 by the following table

Table 2.2. Fortran statement for a normal model

```
      S = 21/A
      DO 16   I = 1,N
      X = I/S
      CALL NDTR (X, PX, D)
      DO 15   J = I, N
      IF (I. EQ. J)   GO TO 14
      Z = J/S
      CALL NDTR (Z, P, D)
      PROB (I,J) = (PX - 0.5) . (P - 0.5)
      PROB (J,I) = PROB (I,J)
 14   PROB (I,J) = 0.0
 15   CONTINUE
 16   CONTINUE
```

The computer program could be used to analyze networks of any size N by redimensioning the variable names which are vector arrays, matrix arrays to the size of the network and enter the data card according to the changes. The program prints out the link density and the weak connectivity of the network, as well as the curves of the average number of contacted stations, the probability distribution function of the number of contacted stations at each step and the terminal reliability of the network.

```
C      SIMULATION OF A RANDOM COMMUNICATICN NETWCRK
C
C
C
       DIMENSION IRMAIN(100),NEW(100),ICOPY(100),ZVEC(100),XVEC(100),
      1GVEC(100),RELIAB(100),PROB(100,100),PRDIST(100,100),Y(100),
      1RLIAB(100,100),RNGE(4),XXVEC(3,100)
       REAL LABEL
       DATA IRMAIN/100*1/,NEW/100*0/,ICOPY/100*0/,ZVEC/1.0,99*0.0/,XVEC/
      21.0,99*0.0/,GVEC/100*0.0/,KSTEP/0/,KHECK/0/,KHECK1/0/,SUMND/1.0/,
      3ZNODES/0.0/,KMAX/0/,KKMAX/0/,CONECT/0.0/,DNSITY/0.0/,NDEMIT/0/,
      4LABEL/' K= '/,MC/1/,IX/102530469/,A/0.0/,K/38000/,S/0.0/,Z/0.C/,X/
      50.0/,P/0.0/,PX/0.0/,RNGE/20.0,0.0,21.0,0.0/
       REAL*8 TITLE(12)
       CALL OVFLOW
       READ(5,1) N
     1 FCRMAT(I2)
       WRITE(6,2)
     2 FORMAT('1',T30,'SIMULATION OF A RANDOM COMMUNICATION NETWORK'/' ')
     3 FCRMAT(6A8)
       DC 8 I=1,N
       DO 7 J=1,N
       RLIAB(I,J)=0.0
     7 CONTINUE
     8 CONTINUE
       DC 9  I=1,N
       GVEC(I)=I-1.0
     9 CONTINUE
       DC 10 I=1,3
       XXVEC(I,1)=1.0
    10 CONTINUE
       DO 13 I=1,3
       DO 11 J=2,100
       XXVEC(I,J)=0.0
    11 CCNTINUE
    13 CONTINUE
       DC 400 L=1,3
C
C      INITIALIZE THE MATRIX OF PROBABILITIES BETWEEN STATIONS TO DESIRED
C      VALUES
C
       A=A+1.0
       DC 16 I=1,N
       DC 15 J=1,N
       IF (I.EQ.J) GO TO 14
       PROB(I,J)=A/(N-1)
       GC TO  15
    14 PROB(I,J)=0.0
```

```
   15 CONTINUE
   16 CCNTINUE
      DO 23 I=1,N
      DO 22 J=1,N
      PRDIST(I,J)=0.0
   22 CONTINUE
   23 CONTINUE
      WRITE(6,40)
   40 FORMAT(//,' MATRIX OF PROBABILITIES'//' ')
      DO 50 I=1,N
      WRITE(6,45)(PROB(I,J),J=1,N)
   45 FORMAT(//,1X,21(F4.2))
   50 CONTINUE
C
C     START TO RANGE OVER THE NODES
C
      DC 300 M=1,K
      PRDIST(1,2)=M
      IRMAIN(1)=0
      NEW(1)=1
      ICOPY(1)=1
C
C     START TO RANGE OVER THE NEWLY CONTACTED NODES
C
   73 DC 200 I=1,N
      IF (NEW(I).EQ.0) GO TO 200
C
C     START TO RANGE OVER THE REMAINING NODES
C
      KHECK1=0
      DC 100 J=1,N
      IF(IRMAIN(J).EQ.0) GO TO 100
C
C     GENERATION OF RANDOM NUMBERS BETWEEN 0 AND 1
C
      CALL RANDOM(IX,YFL,1)
      IF(YFL.GT.PROB(I,J)) GO TO 100
      IF (KHECK1.EQ.1) GO TO 74
      NDEMIT=NDEMIT+1
      KHECK1=1
   74 IF (KHECK.EQ.1) GO TO 80
      DC 75 I1=1,N
      ICCPY(I1)=0
   75 CONTINUE
      KHECK=1
   80 IRMAIN(J)=0
      SUMND=SUMND+1.0
      ICCPY(J)=1
```

```
      ZNODES=ZNODES+1.0
C
C     TEST IF THERE REMAINS ANY NODE
C
      DO 90 I2=1,N
      IF(IRMAIN(I2).EQ.0) GO TO 90
      GO TO 100
   90 CONTINUE
      KSTEP=KSTEP+1
      IF(KMAX.GT.KSTEP) GO TO 220
      KMAX=KSTEP
   95 GO TO 220
  100 CONTINUE
  200 CONTINUE
C
C     TEST TO SEE IF THERE IS ANY NEWLY CONTACTED NODES
C
      DO 210 I3=1,N
      IF (NEW(I3).NE.ICOPY(I3)) GO TO 215
  210 CONTINUE
      GO TO 225
  215 DO 216 I4=1,N
      NEW(I4)=ICOPY(I4)
  216 CONTINUE
  218 KSTEP=KSTEP+1
      IF(KMAX.GT.KSTEP) GO TO 219
      KMAX=KSTEP
  219 KHECK=0
      PRDIST(KSTEP+1,ZNODES+1.0)=PRDIST(KSTEP+1,ZNODES+1.0)+1.0
      ZVEC(KSTEP+1)=ZVEC(KSTEP+1)+ZNODES
      ZNODES=0.0
      GO TO 73
  220 ZVEC(KSTEP+1)=ZVEC(KSTEP+1)+ZNODES
      PRDIST(KSTEP+1,ZNODES+1.0)=PRDIST(KSTEP+1,ZNODES+1.0)+1.0
      ZNODES=0.0
  225 KKMAX=KSTEP+2
      DO 228 I6=KKMAX,N
      PRDIST(I6,1)=PRDIST(I6,1)+1.0
  228 CONTINUE
      IF (NDEMIT.EQ.0) GO TO 230
      DNSITY=(SUMND-1.0)/NDEMIT+DNSITY
      NDEMIT=0
  230 KSTEP=0
      KHECK=0
      SUMND=1.0
  245 DO 247 I=1,N
      IRMAIN(I)=1
      NEW(I)=0
```

```
      ICOPY(I)=0
  247 CONTINUE
  300 CCNTINUE
C
C
C     END OF THE SAMPLING. BEGIN TO COMPUTE  THE DESIRED QUANTITIES
C
C
      DC 310 I=2,N
      ZVEC(I)=ZVEC(I)/K
  310 CONTINUE
      DNSITY=DNSITY/K
      KMAX=KMAX+1
      DO 311 I=2,KMAX
      XXVEC(L,I)=XXVEC(L,I-1)+ZVEC(I)
  311 CONTINUE
C
C     PLCTTING OF THE AVERAGE NUMBER OF CONTACTED NODES AT EACH STEP
C
      CALL UTPLOT(GVEC,ZVEC,KMAX,RNGE,1,0)
      WRITE(6,316)(ZVEC(I),I=1,15)
  316 FORMAT(//,' AVERAGE NUMBER OF NEW NODES AT EACH STEP'//,' (',15(F7
     2.4,1X),')'//' ')
      DC 318 J=1,N
      RLIAB(L,J)=ZVEC(J)/(N-1)
  318 CCNTINUE
      CONECT=XXVEC(L,KMAX)/N
      WRITE(6,319)(RLIAB(L,J),J=1,N),CONECT,DNSITY
  319 FORMAT(//,' TERMINAL RELIABILITY'//,' (',21(F5.2,1X),')',//,' WEAK
     3CCNNECTIVITY = ',F7.5,//,' LINK DENSITY = ',F7.5)
      DNSITY=0.0
      DC 328 I=1,N
      DO 327 J=1,N
      PRDIST(I,J)=PRDIST(I,J)/K
  327 CONTINUE
  328 CONTINUE
      DC 330 I=1,N
      WRITE(6,329)(PRDIST(I,J),J=1,N)
  329 FORMAT(//,1X,21(F5.3,1X))
  330 CCNTINUE
      WRITE(6,331)
  331 FORMAT(//,' PROBABILITY DISTRIBUTION OFTHE NUMBER OF NEW NODES'//
     1' ')
      DO 332 I=2,N
      ZVEC(I)=0.0
      RELIAB(I)=0.0
  332 CONTINUE
      CCNECT=0.0
```

```
C
C     PLOTTING OF THE PROBABILITY DISTRIBUTION OF THE NUMBER OF NEW
C     NODES AT EACH STEP
C
      READ(5,3)TITLE
      DO 347 I=2,4
      DO 343 J=1,N
      Y(J)=PRDIST(I,J)
  343 CONTINUE
      GO TO (347,346,344,345),I
  344 MC=2
      GO TO 346
  345 MC=3
  346 CALL DRAW(15,GVEC,Y,MC,0,LABEL,TITLE,3.0,0.1,0,0,0,0,5,7,1,LAST)
  347 CONTINUE
      MC=1
  400 CONTINUE
C
C     PLOTTING THE TERMINAL RELIABILITY VERSUS THE STEPS
C
      READ(5,3)TITLE
      DO 449 I=1,3
      DO 444 J=1,N
      RELIAB(J)=RLIAB(I,J)
  444 CONTINUE
      GO TO (448,445,446),I
  445 MC=2
      GO TO 448
  446 MC=3
  448 CALL DRAW(10,GVEC,RELIAB,MC,0,LABEL,TITLE,2.0,0.1,0,0,0,C,5,4,1,
     1LAST)
  449 CONTINUE
      MC=1
C
C     PLOTTING THE AVERAGE NUMBER OF CONTACTED NODES
C
      READ(5,3)TITLE
      DO 459 I=1,3
      DO 454 J=1,KMAX
      XVEC(J)=XXVEC(I,J)
  454 CONTINUE
      GO TO (458,455,456),I
  455 MC=2
      GO TO 458
  456 MC=3
  458 CALL DRAW(12,GVEC,XVEC,MC,0,LABEL,TITLE,2.0,3.0,0,0,0,0,6,7,1,
     2LAST)
  459 CONTINUE
```

```
      DO 462 I=1,3
      WRITE(6,460)(XXVEC(I,J),J=1,15)
  460 FORMAT(//,1X,21(F7.4,1X))
  462 CCNTINUE
      WRITE(6,464)
  464 FORMAT(//,' AVERAGE MESSAGE PROPAGATION ')
      STCP
      END
```

III. COMPARATIVE ANALYSIS

For the purpose of comparison, the computer program has been used to investigate a random communication network having 21 stations whose probability of contact is assumed constant. Three cases have been considered where the probability is equal respectively to 0.05, 0.10, and 0.15. The sampling size is calculated to have an error of 0.01 and the results are compared with the theoretical values.

A. THE CALCUATION OF THE SAMPLING SIZE

Suppose that one station tries to send the message to all other stations. The expected number of contacted stations is given by

$$\mu = (n - 1)p. \tag{3.1}$$

The variance is given by

$$\sigma = (n - 1)pq$$

where
n = the size of the network
p = the probability of contact
$q = 1 - p$.

Equation (3,1) becomes

$$\mu = 1$$

in the case where $n = 21$ and $p = 0.05$.

Let Z(1) designate the random number of contacted stations at the first step in one sample. Since the population is infinite, the mean and the standard deviation of the sampling distribution of means are given by

$$\mu_{Z(1)} = \overline{Z}_{(1)} = \mu$$

and $$\sigma_{Z(1)} = \frac{\sigma}{\sqrt{N}}$$

where N = the sampling size.

For large values of N ($N \geq 30$) the sampling distribution of means is approximately a normal distribution with mean μ and standard deviation $\sigma/\sqrt{N}$. The accuracy of the approximation improves as N gets larger (Central limit theorem).

Suppose that one would like the value of $\overline{Z}(1)$ differ from the population mean μ by 0.01 with a confidence level of 95.45%, one has to satisfy the inequality

$$\left| \overline{Z}(1) - \mu \right| \leq 2\,\sigma_{Z(1)} \leq 0.01 \qquad (3.2)$$

or $$\frac{2\sqrt{(n-1)pq}}{\sqrt{N}} \leq 10^{-2}$$

or $$N \geq 4.10^4 \,.\, (n-1)pq = 4.10^4 \,.\, (21-1)(0.05)(0.95)$$

then $$N \geq 3.8 \times 10^4$$

The program has been run through 3.8×10^4 iterations. The results of the simulation are compared with the theoretical values in the following section.

B. COMPARISON BETWEEN SAMPLING AND THEORETICAL VALUES

a. Average Numbers of Contacted Stations at the First and Second Steps

The average numbers of contacted stations are given by [5]

$$\overline{Z}(1) = (n-1)p$$

$$\overline{Z}(2) = (n\ 1)q \left\{1 - (1-p^2)^{n-2}\right\}.$$

They are calculated for three values of the probability p and compared with the results of the sampling in tables 3.1 and 3.2. The difference between sampling and theoretical results is small.

Table 3.1. Values of $\overline{Z}(1)$.

Probability	Theoretical	Sampling	Difference
0.05	1	0.9978	0.0022
0.10	2	2.0007	-0.0007
0.15	3	2.9870	0.0130

Table 3.2. Values of $\overline{Z}(2)$.

Probability	Theoretical	Sampling	Difference
0.05	0.8825	0.8708	0.0117
0.10	3.1289	3.1334	0.0045
0.15	5.9676	5.9370	0.0306

The sampling size has been calculated for the value of probability equal to 0.05. Table 3.1 shows that the difference between the sampling and theoretical values $\overline{Z}(1) - 1$ is 0.0022, less than 0.01 as it is expected.

b. Average Total Numbers of Contacted Stations

Fig. 3.1 gives the numbers of stations which will be eventually contacted for different values of probability. They are tabulated in Table 3.3 along with the corresponding theoretical values.

Table 3.3. Average total number of contacted stations

Probability	Theoretical	Sampling	Difference
0.05	4.50	4.40	0.10
0.10	14.00	14.07	-0.07
0.15	19.00	19.10	0.10

The theoretical [5] and sampling curves giving the average numbers of contacted stations up to step K are similar. Both show that the propagation of message is over between steps 4 and 6.

c. Terminal Reliability

The terminal reliability curves in [5] and in Fig. 3.2 are very similar. They show that the highest reliability occurs between steps 2 and 3 and is negligible from step 6. The terminal reliability is the measure of the speed of propagation of the message.

Table 3.4. Values of highest terminal reliability

Probability	Theoretical	Sampling	Difference
0.10	0.175	0.175	0.0
0.15	0.34	0.32	0.02

d. Probability Distribution of the Number of New Stations Contacted at Each Step

Fig. 3.3, 3.4, and 3.5 give the probability distribution of new stations contacted at each step, they correspond to the formula (1.3), they are very similar to the theoretical curves in [5].

In summary, the computer program has yielded results almost identical to the theoretical values if the sampling size is large, between 4 x 10^4 and 10^5. Since the program was written without any constraint, any assumption on the probability of communication between pairs of stations, it could be used with confidence to study any random network on the condition that the matrix of probabilities is initialized to values which reflect each particular circumstance. As an example, in the next chapter, the program is used to investigate a communication network whose destruction is supposed normally distributed around a nuclear explosion or an earthquake.

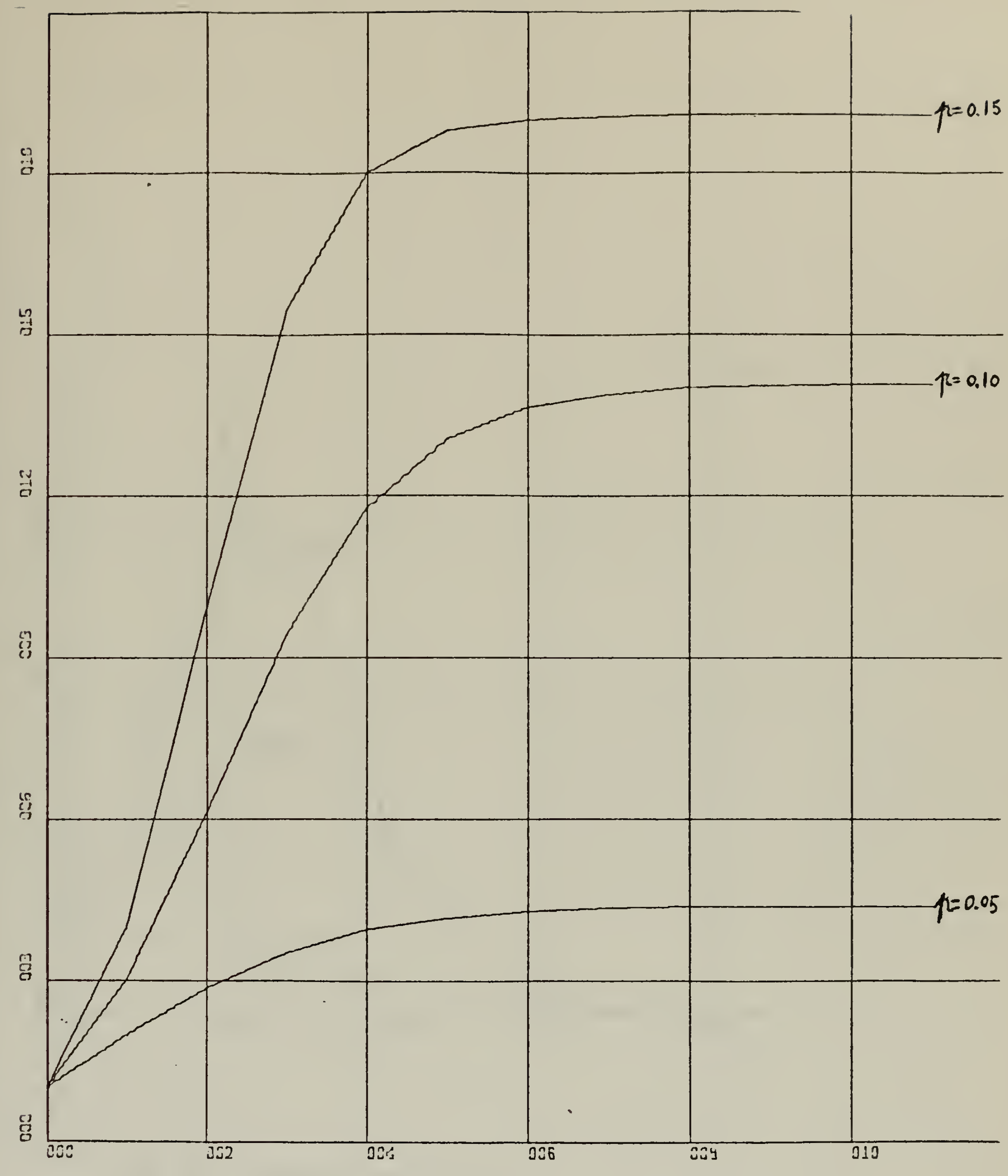


Fig. 3.1. Average message propagation

x scale = order of step (2 units per inch)

y scale = number of contacted stations (3 units per inch)

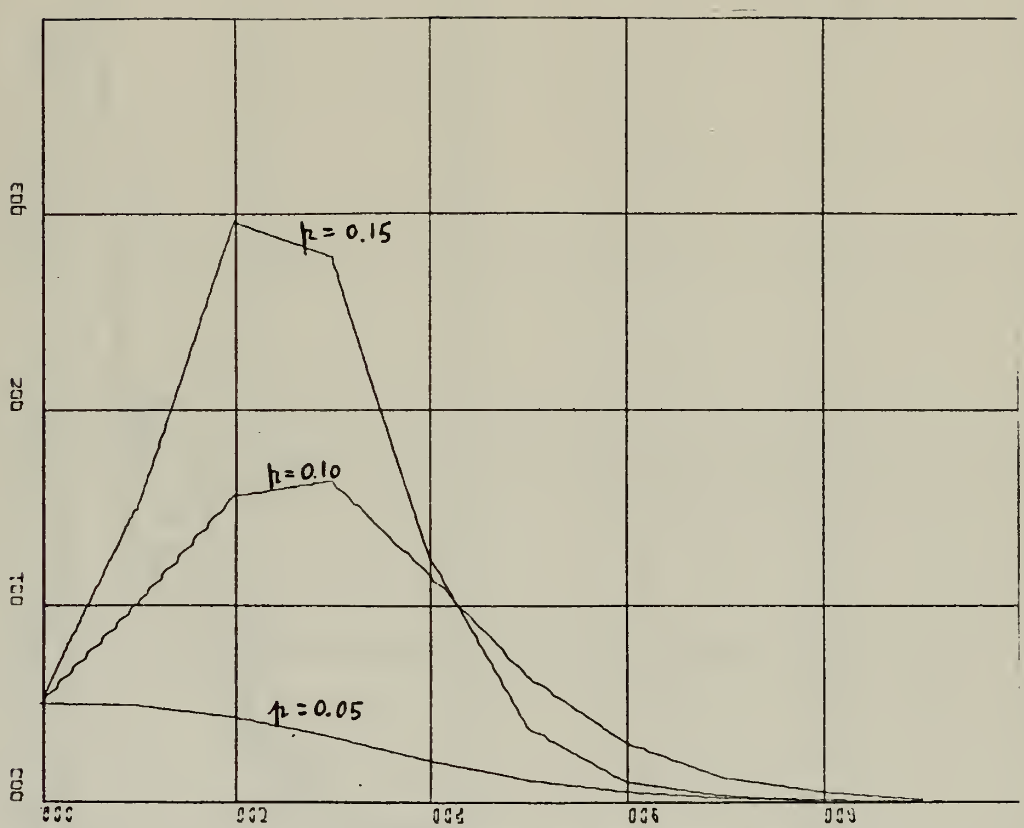


Fig. 3.2. Terminal Reliability

x scale = order of step (2 units per inch)

y scale = terminal reliability (0.1 unit per inch)

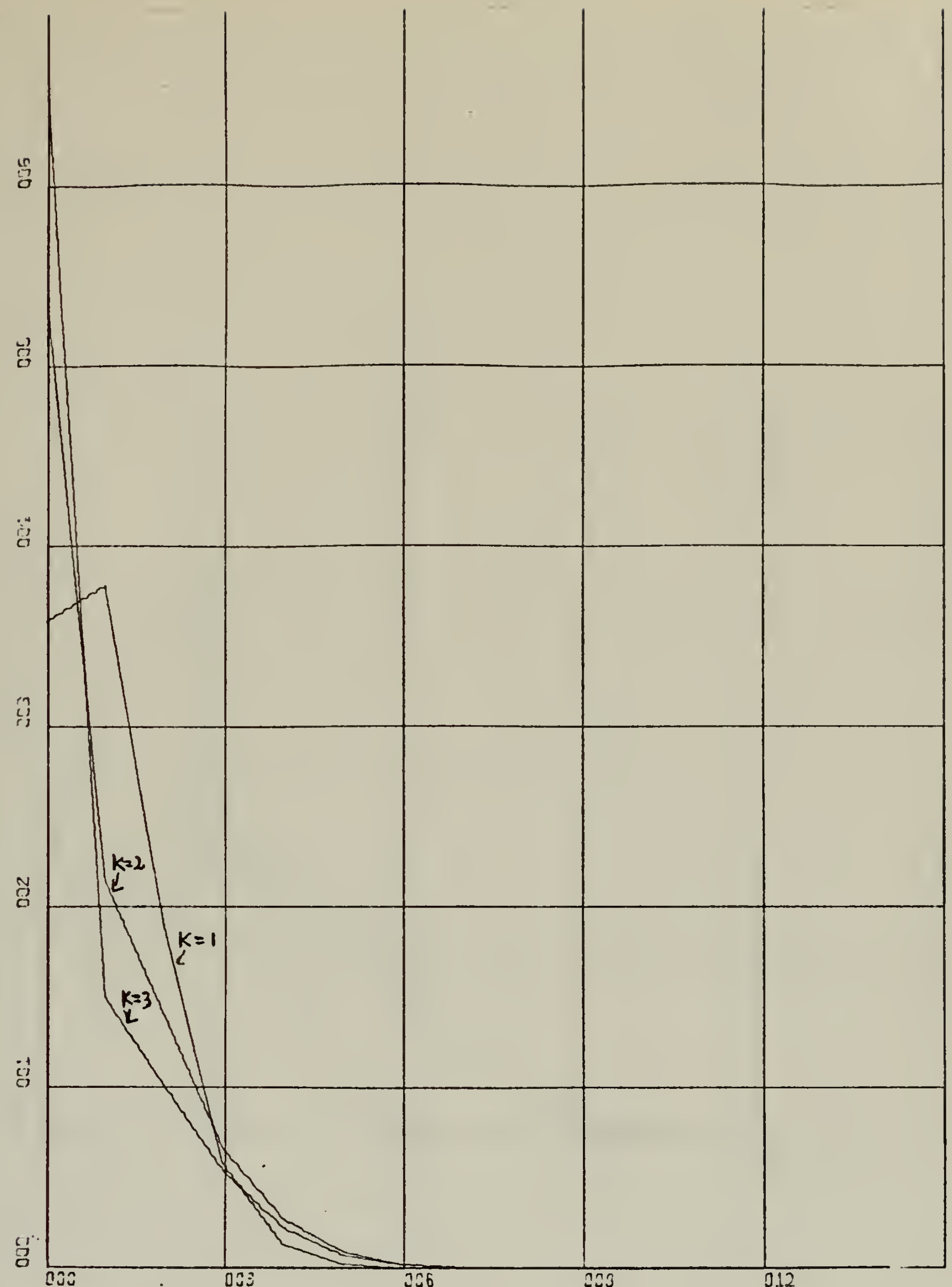


Fig. 3.3. Probability distribution of the number of contacted stations at each step.

Probability = 0.05

x scale = number of contacted stations (3 units per inch)

y scale = probability distribution (0.1 unit per inch)

k = order of step

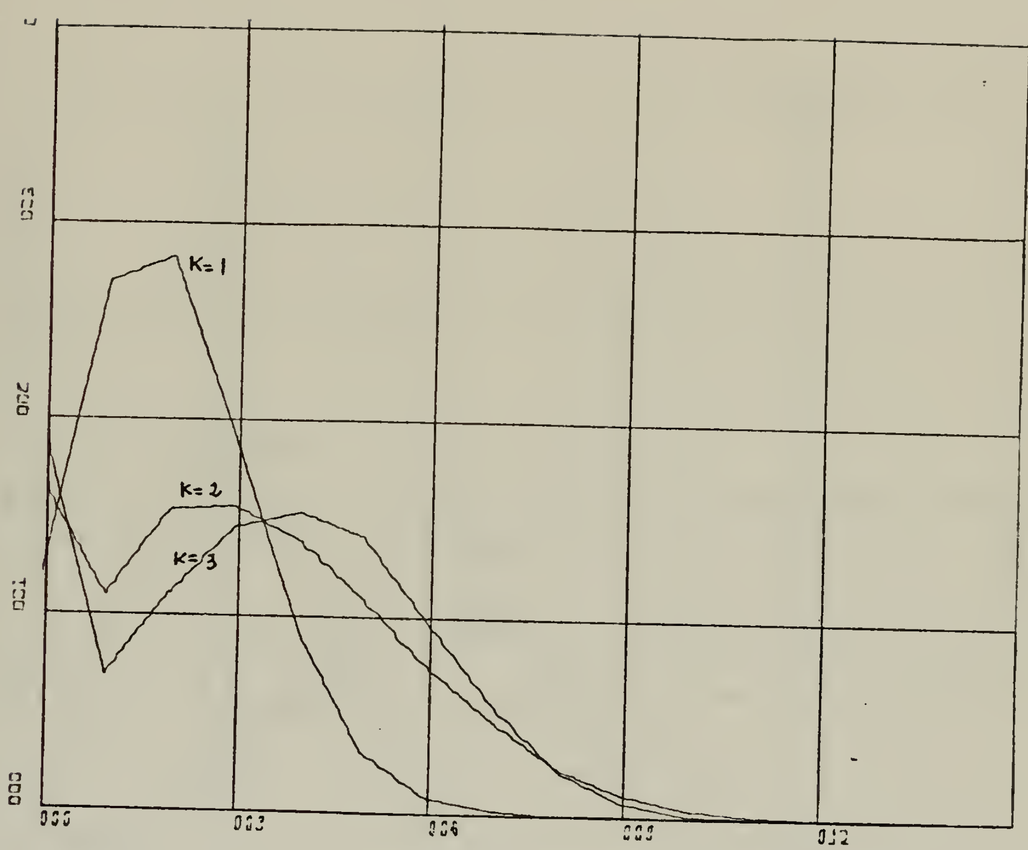


Fig. 3.4. Probability distribution of the number of contacted stations at each step.

Probability = 0.1

x scale = number of contacted stations (3 units per inch)

y scale = probability distribution (0.1 unit per inch)

k = order of step

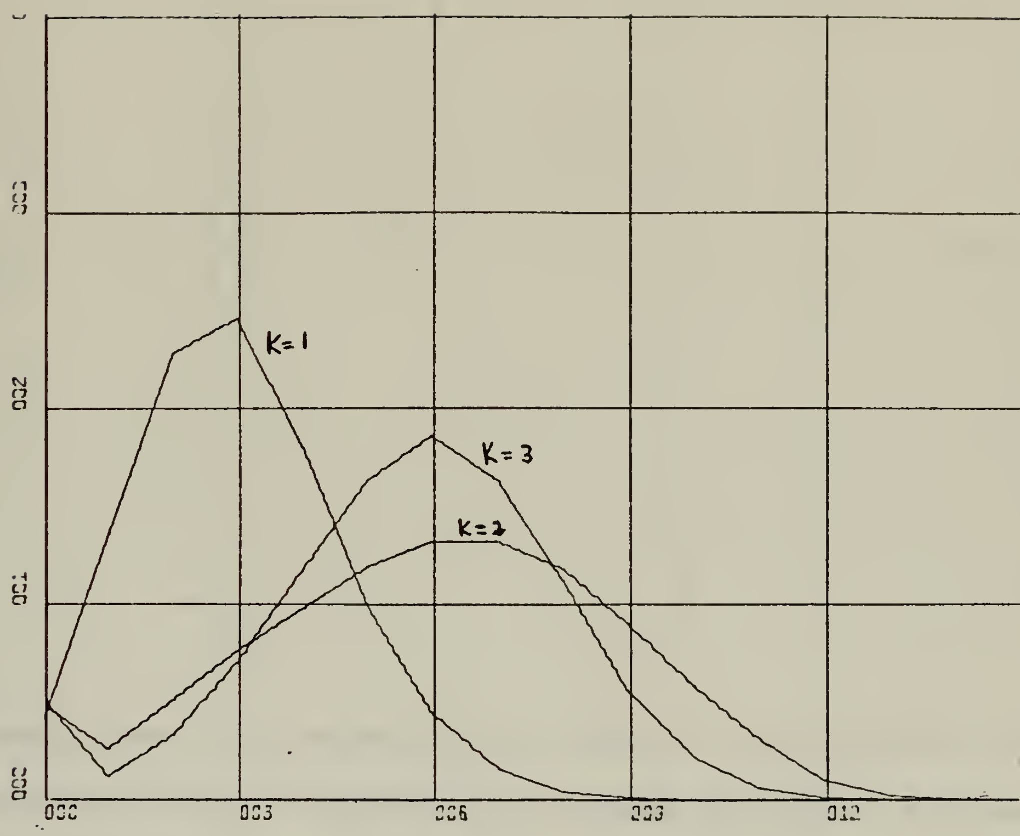


Fig. 3.5. Probability distribution of the numbers of contacted stations at each step.

Probability = 0.15

x scale = number of contacted stations (3 units per inch)

y scale = probability distribution (0.1 unit per inch)

k = order of step

IV. INVESTIGATION OF A GENERAL MODEL

A. REPRESENTATION OF THE MODEL

Consider a communication network of n stations which may be represented by an n x n matrix whose elements are 1 or 0 as

$$\underset{\sim}{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ . \\ . \\ . \\ . \\ . \\ . \\ n \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & 2 & 3 & . & . & . & . & . & n & \end{array} \\ \left(\begin{array}{cccccccccc} 0 & 1 & 0 & 0 & 1 & . & . & . & 1 \\ 0 & 0 & 1 & 0 & 0 & . & . & . & 0 \\ 1 & 0 & 0 & 0 & . & . & . & . & 1 \\ 0 & . & . & . & . & . & . & . & . \\ 1 & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \end{array}\right) \end{array} \qquad (4.1)$$

Suppose there is a perturbation near station 1, whose effect is to destroy the links of communication according to a normal distribution $N(0,\sigma)$ around a center. This means that the destruction is maximum at the point of impact and decreases radially according to the equation

$$p(x) = \int_0^x \frac{1}{\sigma\sqrt{2\pi}} \; e^{-1/2t^2/\sigma^2} \, dt.$$

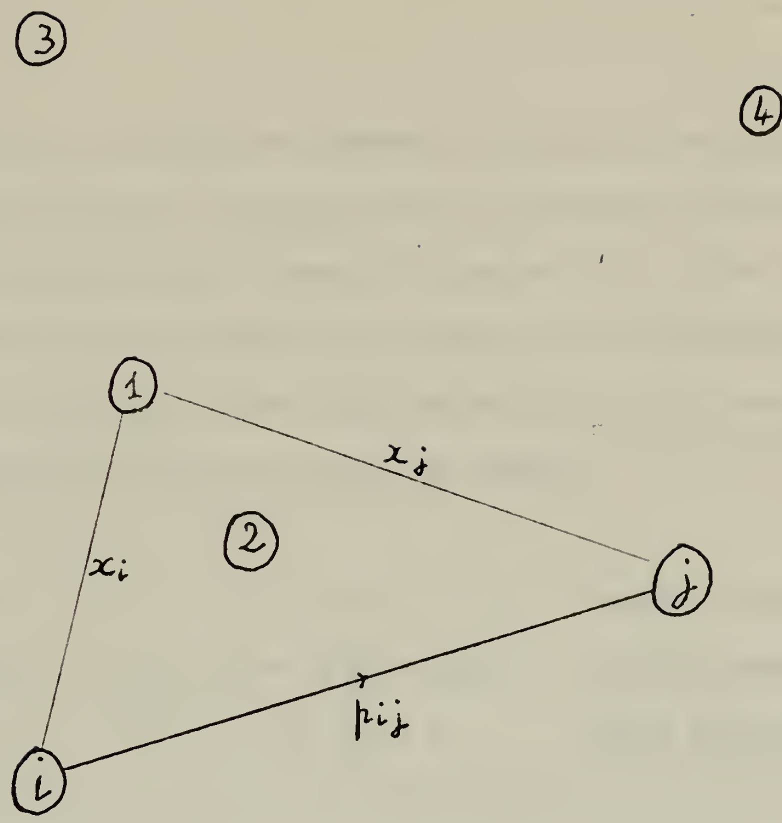


Fig. 4.1. Illustrating the equation (4.2)

where p(x) = probability of survival of a link of a station

x = distance of a station to the impact

σ = standard deviation of the destruction

The probability of communication between stations i and j is

$$p_{ij} = p(x_i)\ p(x_j)$$

$$= \int_0^{x_i} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} dt . \int_0^{x_j} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} dt. \qquad (4.2)$$

Suppose that the stations are numbered according to increasing distance and the network is complete, then all elements of the matrix $\underset{\sim}{P}$ have to be replaced by the values of equation (4.2). The value of the standard deviation measures the extent of the destruction. The simulation was made for three cases where station 21 is one, two and three standard deviations away from the impact.

first case	$\sigma = 21$	extensive destruction
second case	$\sigma = \frac{21}{2} = 10.5$	medium destruction
third case	$\sigma = \frac{21}{3} = 7$	minor destruction

The subscripts of the stations which represent their distance to the impact, have been expressed in standard units and the subroutine NDTR has been used to initialize the matrix $\underset{\sim}{P}$. The sampling size has been taken equal to 3.8×10^4. The results are summarized in the following section.

B. RESULTS OF THE SIMULATION

When the farthest station is one standard deviation from the impact, i.e, the case when the destruction is extensive, the message is not carried to many stations. On the average, 1.26 stations receive the message, the link probabilities being very low. In the case of medium destruction, the farthest station two standard deviations from the impact, 4.17 stations receive the message. The propagation

ceases at step 5. 7.45 stations receive the message when station 21 is three standard deviations away from the impact. The above results are obtained with the assumption the communication network forms a complete graph, i.e., there is a link between any pair of stations. It is usually not the case in real life, therefore the above results are only the upper bound of number of stations which will receive the message in case of a war.

If one compares with the model where the link probability is constant, one finds the latter model always gives much higher results than the current model which is nearer to reality. Table 4.1 summarizes the results of two models.

Table 4.1. Number of contacted stations in two models compared.

Probability	Mattei	Standard Deviation	Normal Model
0.05	4.40	21	1.25
0.10	14.07	10.5	4.17
0.15	19.10	7	7.45

Fig. 4.2 shows the average message propagation in the three cases. The weak connectivity of the normal model is also smaller than the Mattei model, it is equal to 0.06, 0.20 and 0.35 when the farthest station is away from the impact one, two and three standard deviations respectively. The link densities are respectively 0.08, 0.36 and 0.67. All results point out that in the normal model, the propagation stops rapidly. The terminal reliability in Fig. 4.3 is a monotonically decreasing function in the first case, attains the maximum

value at step 3 for the second and third cases. It means that the propagation increases, attains the culmination then decreases and stops in average at step 5. The curves of probability distribution function in Figs. 4.4, 4.5 and 4.6 of the number of contacted stations at each step show also small values compared with the Mattei model.

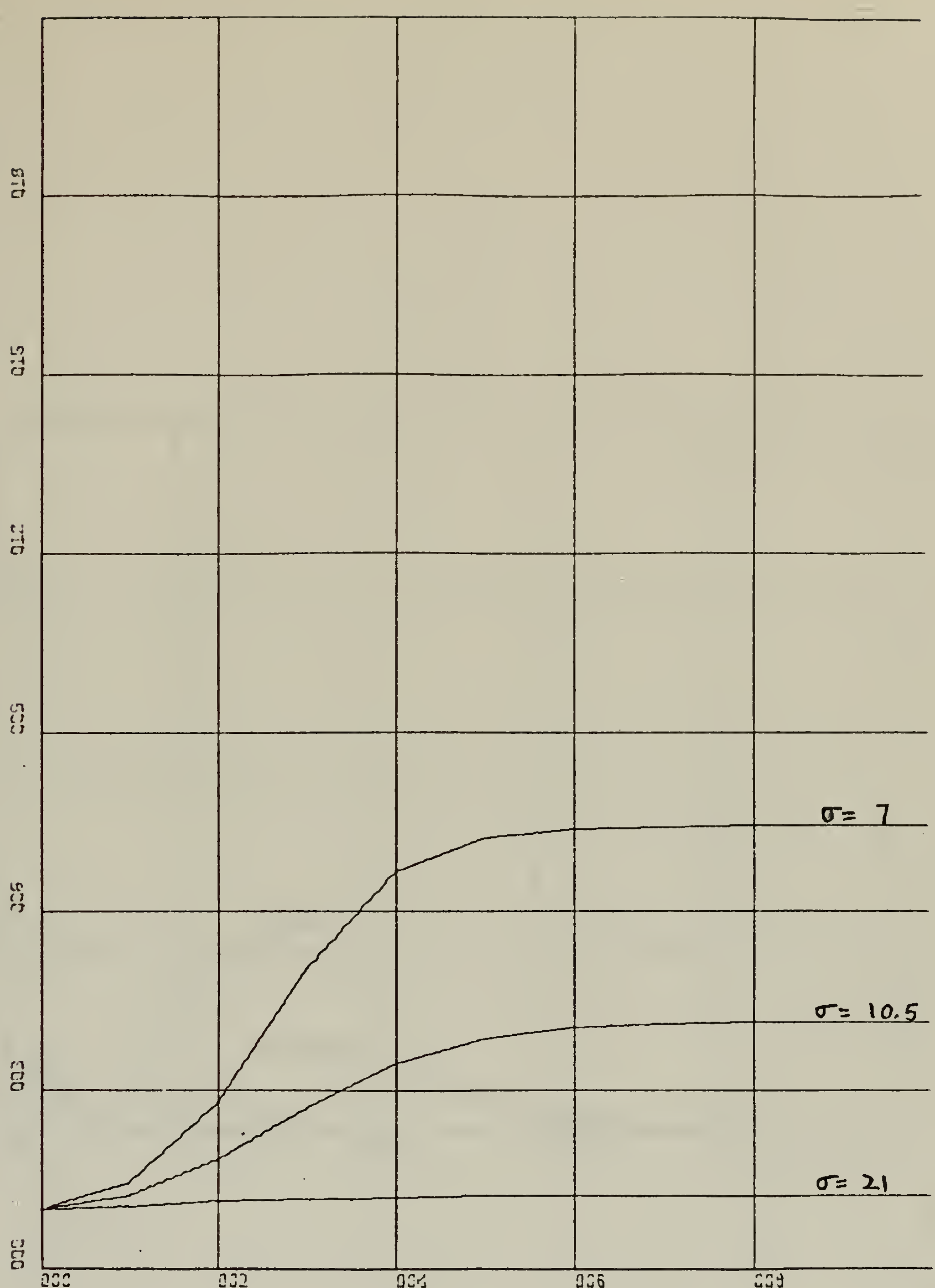


Fig. 4.2. Average message propagation

x scale = order of step (2 units per inch)

y scale = number of contacted stations (3 units per inch)

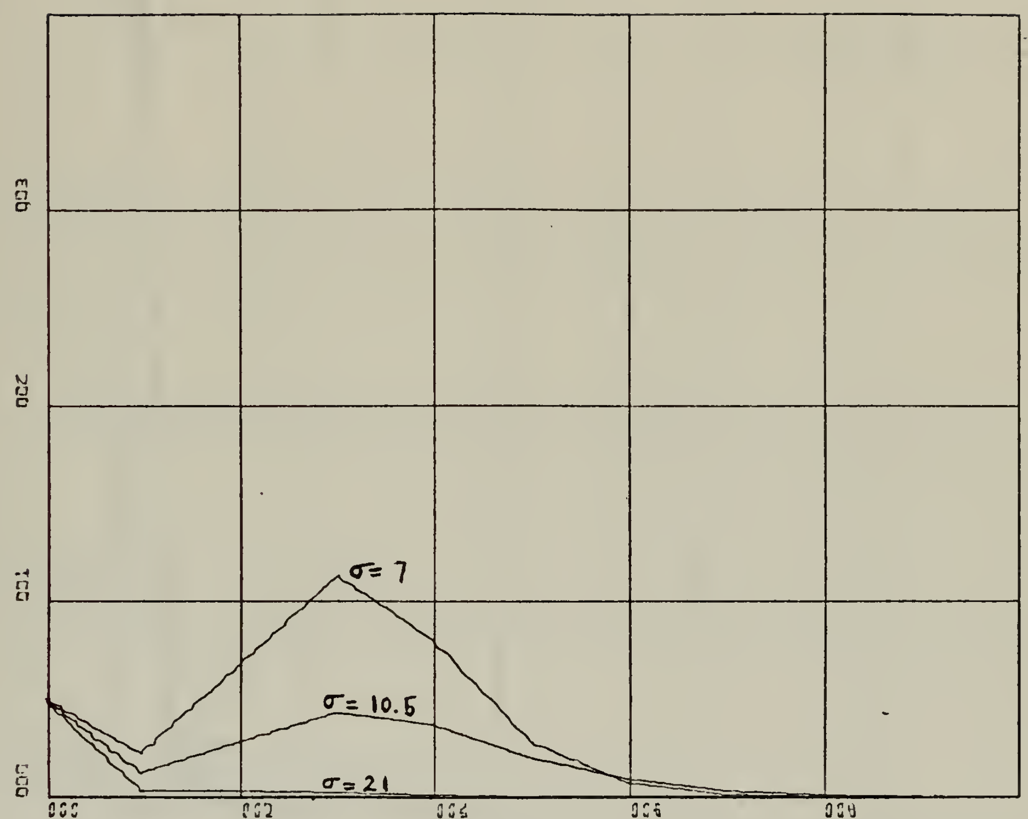


Fig. 4.3. Terminal reliability

x scale = order of step (1 unit per inch)

y scale = terminal reliability (0.1 unit per inch)

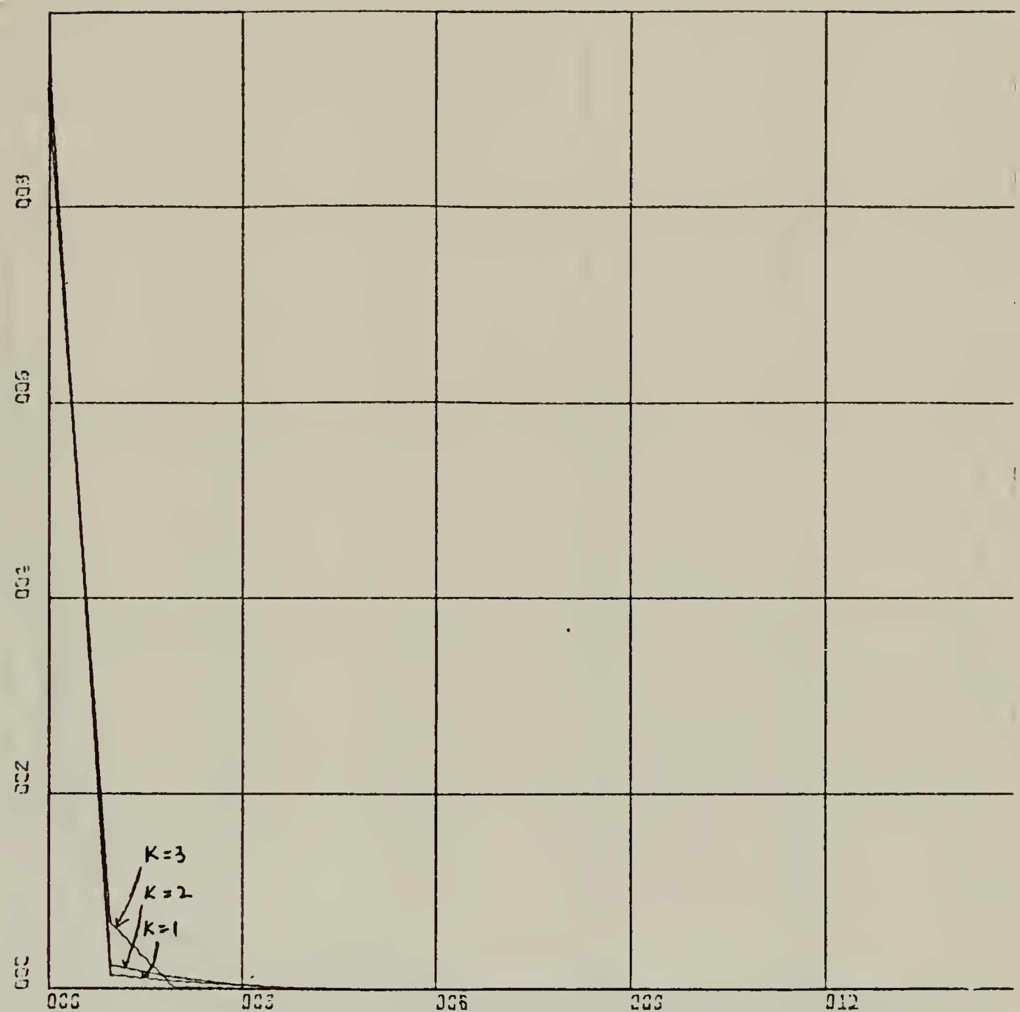


Fig. 4.4. Probability distribution of the number of contacted stations at each step

Standard deviation = 21

x scale = number of contacted stations (3 units per inch)

y scale = probability distribution (0.2 unit per inch)

k = order of step

Fig. 4.5. Probability distribution of the number of contacted stations at each step

Standard deviation = 10.5

x scale = number of contacted stations (3 units per inch)

y scale = probability distribution (0.2 unit per inch)

k = order of step

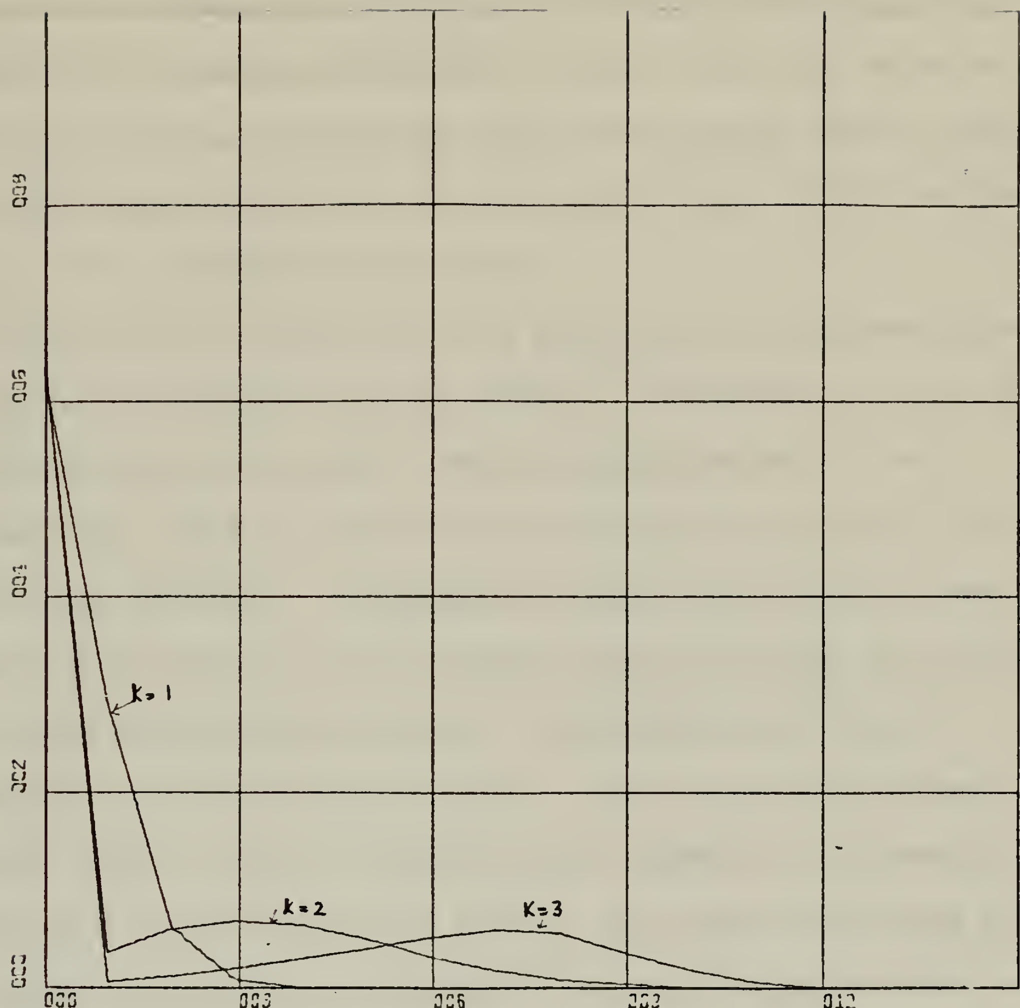


Fig. 4.6. Probability distribution of the number of contacted stations at each step

Standard deviation = 7

x scale = number of contacted stations (3 units per inch)

y scale = probability distribution (0.2 unit per inch)

k = order of step

V. CONCLUSION

The analysis of the message propagation through a random communication network has been studied by various investigators. To be able to solve the problem, each has adopted different assumptions on the probability of communication between stations. The main effort of this paper has been to solve the same problem without having to make any simplifying assumption on the link probability. The latter could vary to fit a particular circumstance.

A flowchart has been drawn and a program for its implementation on the digital computer has been written. The program gives the desired quantities of a random communication network whose link probabilities could be varied to fit a particular situation by the Monte Carlo technique. The program has been tested with the model whose link probabilities were assumed constant and whose characteristics could be derived analytically. The results were found in agreement with the analytical results. The program has also been used on a model where the destruction was supposed to be normally distributed around the point of attack. The results also agree with the intuitive idea that, as the effectiveness of the destruction of the network increases, the propagation of message stops rapidly.

In the computer program, the time taken by the message to travel between any pair of stations is assumed constant, which is not true in real life. Some stations take longer time to process and send the information than the others. It is proposed for future work that the time t_{ij}, time taken by station i to process information and send it to station j, could be varied at will. With this second

generalization added, the simulation will give the average time of the propagation of the message and what routing procedure will give the fastest communication between various stations.

LIST OF REFERENCES

1. Bailey, N. T. J., "A Simple Stochastic Epidemic," Biometrika, 37, pp. 193-202.

2. Landau, H. G. and Rapoport, A., "Contribution to the Mathematical Theory of Contagion and Spread of Information: I, Spread Through a Thoroughly Mixed Population," Bull. Math. Biophys., Vol. 15, 1953.

3. Landahl, H. D., "On the Spread of Information with Time and Distance," Bull. Math. Biophys., Vol. 15, 1953.

4. Prihar, Z., "Communication Nets in Random Behavior," IEEE Trans. on Com. Tech., Vol. 14, pp. 389-399, August 1966.

5. Mattei, A., "Message Propagation Through Random Communication Networks," Drexel Institute of Technology, Ph. D., 1971.

6. Shreider, Y. A., "The Monte Carlo Method," Pergamon Press, 1966.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Dr. Sydney Parker, Department Chairman, Code 52 Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1
4.	Professor Shu-Gar Chan Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	2
5.	Vietnamese Navy Headquarters KBC 3317 Saigon, Vietnam	2
6.	University of Saigon Faculty of Sciences Saigon, Vietnam	2
7.	Trung tam quoc gia Ky Thuat Pho thu Saigon, Vietnam	2
8.	Cdr. Nguyen dịch Hung, VNN 174/107/8 Nguyen thein Thuat Saigon, Vietnam	2

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION 2b. GROUP

3. REPORT TITLE

Analysis of a Random Communication Network by Simulation

4. DESCRIPTIVE NOTES *(Type of report and, inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*

Nguyen dich Hung

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1973	66	6

8a. CONTRACT OR GRANT NO. b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

In this paper a unified presentation is made on the results of various investigators on the properties of random communication networks. These results are interpreted in such a way that the properties may be determined by using a digital computer with the application of the Monte Carlo method. The computer program is written and tested. Results for some networks are compared with theoretical values.

Security Classification

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Random Communication Network						
Digital Computer						
Monte Carlo						

Thesis 145276
H925 Hung
c.1 Analysis of a random
communication network by
stimulation.